

日本国特許
JAPAN PATENT OFFICE

D. J 11-16-01
A2
Priority
Paper

J1040 U.S. PTO
09/973774



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2000年10月12日

出願番号

Application Number:

特願2000-312476

出願人

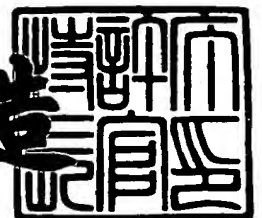
Applicant(s):

パイオニア株式会社

2001年 8月 3日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2001-3070027

【書類名】 特許願

【整理番号】 54P0634

【提出日】 平成12年10月12日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00
G08G 1/0969

【発明者】

【住所又は居所】 埼玉県川越市山田字西町 2 5 番地 1 パイオニア株式会
社川越工場内

【氏名】 斎藤 宏

【特許出願人】

【識別番号】 000005016

【氏名又は名称】 パイオニア株式会社

【代理人】

【識別番号】 100063565

【弁理士】

【氏名又は名称】 小橋 信淳

【手数料の表示】

【予納台帳番号】 011659

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声認識装置ならびに方法

【特許請求の範囲】

【請求項 1】 複数の基準音声信号を有する複数の音声認識辞書を互いに階層的に関連付けて記憶した階層辞書部と、前記階層辞書部から適切な音声認識辞書を抽出する抽出手段と、前記抽出された音声認識辞書を記憶する一覧記憶手段と、音声を入力する音声入力手段と、入力された音声と前記一覧記憶手段に記憶されている音声認識辞書内の基準音声情報とを比較することによって入力された前記音声を認識する認識手段と、認識された前記音声に相当する基準音声情報の下位階層に属する音声認識辞書を抽出手段が抽出し、抽出された音声認識辞書を前記一覧記憶手段が更新記憶する音声認識装置であって、

所定の音声認識辞書に階層飛ばしを示す基準音声情報が用意されており、前記認識手段が階層飛ばしを示す基準音声情報に相当する音声の入力を認識した際に、前記一覧記憶手段に記憶されていた基準音声情報の下位階層に属する音声認識辞書を前記抽出手段が抽出し、前記一覧記憶手段に更新記憶することを特徴する音声認識装置。

【請求項 2】 前記音声認識辞書として、施設の種別名を記憶した種別辞書と、前記施設の種別に属する施設名を記憶した地点辞書とを備えることを特徴とする請求項 1 に記載の音声認識装置。

【請求項 3】 前記音声認識辞書として、地域名を記憶した地域辞書と、いずれかの地域に存在する施設の施設名を記憶した地点辞書とを備えることを特徴とする請求項 1 に記載の音声認識装置。

【請求項 4】 前記音声認識辞書として、地域名を記憶した地域辞書と、前記施設の種別名を記憶した種別辞書と、いずれかの地域に存在するとともにいずれかの種別に属する施設の施設名を記憶した地点辞書とを備え、

前記種別名選択階層にて階層飛ばしを示す基準音声情報が認識された後、前記抽出手段が前記地域辞書を抽出することを特徴とする請求項 1 に記載の音声認識装置。

【請求項 5】 前記認識手段によって認識された前記基準音声情報の下位階

層に属する音声認識辞書の中の基準音声情報の数を判定する数判断手段と、入力された音声記憶する入力音声記憶手段と、前記入力音声記憶手段に記憶されている音声と一覽記憶手段に記憶されている基準音声情報とを前記認識手段が順次比較することにより、類似する基準音声情報を認識し、類似する基準音声情報を記憶する類似ワード記憶手段とを備えた音声認識装置であって、

前記数判断手段において、認識された音声に相当する基準音声情報の下位階層に属する音声認識辞書内の基準音声情報の語句の数が基準値を超えているか判断する判断手段を備え、

前記所定数以上と判断されたとき、前記抽出手段が下位階層に属する前記音声認識辞書の一部の音声認識辞書を抽出して前記一覽記憶手段に記憶し、

前記認識手段が前記一覽記憶手段に記憶されている基準音声情報との比較を終了した後、前記抽出手段が下位階層に属する前記音声認識辞書のうち未抽出の辞書を抽出し、前記一覽記憶手段が更新記憶するとともに、

前記認識手段が前記一覽記憶手段内に更新記憶された辞書に属する基準音声情報と前記入力音声記憶手段に記憶されている音声とを順次比較し、類似する基準音声情報を認識し、

類似ワード記憶手段が新たに認識された類似する前記基準音声情報を追加記憶することを特徴とする請求項 1 に記載の音声認識装置。

【請求項 6】 前記類似ワード記憶手段に記憶された全ての類似ワードにおける一つを前記認識手段が認識し、認識結果とすることを特徴とする請求項 5 に記載の音声認識装置。

【請求項 7】 前記一覽記憶手段に記憶されている基準音声情報の内、複数の類似する基準音声情報を前記類似ワード記憶手段に記憶し、前記類似ワード記憶手段に記憶された全ての類似する基準音声情報のなかから更に認識結果を選出する選出手段を備えることを特徴とする請求項 5 に記載の音声認識装置。

【請求項 8】 複数の基準音声情報を有する複数の音声認識辞書を記憶した階層辞書部と、前記複数の音声認識辞書のうち 1 の辞書を抽出する抽出手段と、抽出された前記辞書を記憶する一覽記憶手段と、音声を入力する音声入力手段と、入力された音声を記憶する入力音声記憶手段と、前記入力音声記憶手段に記憶

されている音声と前記一覧記憶手段に記憶されている基準音声情報と順次比較し、類似する基準音声情報を認識する認識手段と、類似する前記基準音声情報を記憶する類似ワード記憶手段とを備えた音声認識装置であって、

前記一覧記憶手段に記憶されている辞書に属する全ての基準音声情報と前記入力音声記憶手段に記憶されている音声との比較を前記認識手段が終了した後、前記抽出手段が音声認識辞書から未抽出の辞書を抽出し、前記一覧記憶手段が更新記憶するとともに、

前記認識手段が前記一覧記憶手段に更新記憶された辞書に属する基準音声情報と前記入力音声記憶手段に記憶されている音声との比較をして類似する基準音声情報を認識し、

前記類似ワード記憶手段が新たに認識した類似する基準音声情報を追加記憶することを特徴とする音声認識装置。

【請求項 9】 前記類似ワード記憶手段に記憶されている複数の基準音声情報の中から、更に認識結果を選出する選出手段を備えたことを特徴とする請求項 8 に記載の音声認識装置。

【請求項 10】 階層構造をなす複数の音声認識辞書の中から基準音声情報を抽出し、抽出された基準音声情報と入力された音声とを比較することにより音声認識する音声認識方法であって、

所定の音声認識辞書に階層飛ばしを示す基準音声情報が用意されており、前記階層飛ばしを示す基準音声情報に相当する音声の入力を認識した際に、比較対象となっていた基準音声情報の下位階層に属する音声認識辞書の一部を抽出して音声認識を行うことを特徴とする音声認識方法。

【請求項 11】 認識された基準音声情報の下位階層に属する音声認識辞書の中の基準音声情報の数を判断し、その数が基準値を超えていると判断されたとき、下位階層に属する前記音声認識辞書の一部を抽出して、比較することにより類似する基準音声情報を認識し、抽出された基準音声情報との比較が終了した後、

下位階層に属する前記音声認識辞書の中から未抽出の音声認識辞書を抽出し、比較することにより類似する基準音声情報を認識し、

複数の類似する前記基準音声情報の中から、入力された音声に相当する基準音声情報をさらに選出することを特徴とする請求項 1 0 に記載の音声認識方法。

【請求項 1 2】 複数の基準音声情報を有する複数の音声認識辞書の中から 1 つの音声認識辞書を抽出し、抽出された音声認識辞書内の基準音声情報と入力された音声を比較し、抽出された前記音声認識辞書による基準音声情報との比較が終了した後、前記 1 つの音声認識辞書と異なる他の音声認識辞書を抽出し、抽出された音声認識辞書内の基準音声情報を、比較対象の基準音声情報として更新すると共に、更新された基準音声情報と前記入力された音声とを比較することにより、入力された前記音声を認識することを特徴とする音声認識方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、ユーザにより入力された音声を認識して機器の制御を行う音声認識装置ならびに音声認識方法に関するもので、特に音声入力操作の改善をはかるものである。

【0 0 0 2】

【従来の技術】

所望の地点を指定して目的地として設定し、車両の現在位置から目的地までの経路を探索し、これをその現在位置を含む地図と共にディスプレイ上に表示させることにより、目的地までの誘導を行うカーナビゲーションとして、近年、マイクを用いて、入力されたユーザの音声を認識し、認識された音声によって各種操作を行えるようにした音声操作機能付きのものが登場してきた。

【0 0 0 3】

カーナビゲーションにおける音声操作による地点の指定は、学校、病院、駅といった対照地点に存在する施設の種別や、地点の住所などを、ガイダンスメッセージに応じて順次音声を発話して入力し、最終的に地点の固有名称を発話することにより行われている。このように、発話毎の認識対象語句を設定し、続く認識対象語句の絞込みを行うことにより、認識率の確保を図っている。

【0 0 0 4】

目的地設定手順の一例を説明する。初期の段階では、音声認識システム側では、カーナビゲーションを操作するための制御コマンド認識対象語句とした”制御コマンド辞書”が設定されており、ユーザが「目的地設定」とコマンドを発話することでシステムに目的地までの経路設定を行う意思を伝える。

続いて目的地となる具体的な地点を指定する必要があるが、カーナビゲーションに用意されている地点が膨大であるため、一度の音声操作で指定することは認識率やメモリサイズの制約から実用的でない。そこで、指定対象となる地点の数の絞込みを行う。

【0005】

はじめに対象地点に存在する施設の種別（以下ジャンルと称する）での絞込みを行う。音声認識辞書を”制御コマンド辞書”から”ジャンル名称辞書”に入れ替え、(1)「ジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してジャンルの発話を促す。これに対してユーザがジャンルとして例えば(2)「教育施設」と発話すると音声認識システムが発話を認識する。更なる絞込みを行うため教育施設に属する更に詳細なサブジャンルを指定するように、音声認識辞書を”ジャンル名称辞書”から”教育施設に属するサブジャンル名称辞書”に入れ替え、(3)「次のジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してサブジャンル名称の発話を促す。これに対してユーザがサブジャンルとして例えば(4)「大学」と発話すると音声認識システムが発話を認識する。

【0006】

サブジャンルが確定すると、更に、地域での絞込みを行う。音声認識辞書を”サブジャンル名称辞書”から”都道府県名称辞書”に入れ替え、(5)「都道府県名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して都道府県名の発話を促す。これに対してユーザが、例えば(6)「東京都」と発話すると音声認識システムが発話を「東京都」と認識する。サブジャンルが「大学」でかつ都道府県名が「東京都」だった場合、更に詳細に市区町村名を指定するように、あらかじめシステム側で決められている。そのため、音声認識辞書を”都道府県名称辞書”から”東京都の市区町村名辞書”に入れ替え、(7)「市区町村名をどうぞ」とガイ

ダンスメッセージを出力し、ユーザに対して市区町村名の発話を促す。これに対してユーザが例えば(8)「新宿区」と発話すると音声認識システムが発話を認識する。

【0007】

システム側は音声認識辞書を”東京都の市区町村名辞書から”東京都新宿区に存在しかつ大学に属する施設名称を認識対象語句とした”大学名称辞書”に入れ替え、(9)「名称をどうぞ」とガイダンスメッセージを出力し、指定地点の具体的名称の発話を促す。ここで、ユーザが「〇〇大学」と発話すると音声認識システムが認識しナビゲーションが〇〇大学を目的地として設定する。このように対象地点の条件を入力して、対象地点の数を減らして、絞込まれた対象地点の固有名称の入力が行われる。

【0008】

【発明が解決しようとする課題】

ところで、上述した絞込条件や条件の入力の順序は予め決められているため、ユーザが知らない条件の入力が催促されるという事態が発生する。その際にユーザがその催促に回答できなかった場合、以降継続する絞込条件の入力を行うステップへ進むことができなくなってしまう。そのため、目的とする対象地点の具体的名称を発話することなく、地点の指定を断念しなければならないことになり、操作性や応答性の面で難があった。

【0009】

本発明は上記事情に鑑みてなされたものであり、システムから求められた条件の入力を飛ばして、続く条件の入力へ進むための基準音声情報（以下階層飛ばし語句と称する）として「わからない」、「バス」などを用意しておき、ユーザの階層飛ばし語句の入力を受けて、地点の指定を行うための以降の継続するステップに進むことのできる仕組みを構築し、操作性や応答性の改善をはかった音声認識装置ならびに音声認識方法を提供することを目的とする。

【0010】

【課題を解決するための手段】

上述した課題を解決するために請求項1に記載の発明は、複数の基準音声信号

を有する複数の音声認識辞書を互いに階層的に関連付けて記憶した階層辞書部と、前記階層辞書部から適切な音声認識辞書を抽出する抽出手段と、前記抽出された音声認識辞書を記憶する一覧記憶手段と、音声を入力する音声入力手段と、入力された音声と前記一覧記憶手段に記憶されている音声認識辞書内の基準音声情報とを比較することによって入力された前記音声を認識する認識手段と、認識された前記音声に相当する基準音声情報の下位階層に属する音声認識辞書を抽出手段が抽出し、抽出された音声認識辞書を前記一覧記憶手段が更新記憶する音声認識装置であって、所定の音声認識辞書に階層飛ばしを示す基準音声情報が用意されており、前記認識手段が階層飛ばしを示す基準音声情報に相当する音声の入力を認識した際に、前記一覧記憶手段に記憶されていた基準音声情報の下位階層に属する音声認識辞書を前記抽出手段が抽出し、前記一覧記憶手段に更新記憶することを特徴する。

【0011】

請求項2に記載の発明は、請求項1に記載の音声認識装置において、前記音声認識辞書として、施設の種別名を記憶した種別辞書と、前記施設の種別に属する施設名を記憶した地点辞書とを備えることを特徴とする。

【0012】

請求項3に記載の発明は、請求項1に記載の音声認識装置において、前記音声認識辞書として、地域名を記憶した地域辞書と、いずれかの地域に存在する施設の施設名を記憶した地点辞書とを備えることを特徴とする。

【0013】

請求項4に記載の発明は、請求項1に記載の音声認識装置において、前記音声認識辞書として、地域名を記憶した地域辞書と、前記施設の種別名を記憶した種別辞書と、いずれかの地域に存在するとともにいずれかの種別に属する施設の施設名を記憶した地点辞書とを備え、前記種別名選択階層にて階層飛ばしを示す基準音声情報が認識された後、前記抽出手段が前記地域辞書を抽出することを特徴とする。

【0014】

請求項5に記載の発明は、請求項1に記載の音声認識装置において、前記認識

手段によって認識された前記基準音声情報の下位階層に属する音声認識辞書の中の基準音声情報の数を判定する数判断手段と、入力された音声記憶する入力音声記憶手段と、前記入力音声記憶手段に記憶されている音声と一覽記憶手段に記憶されている基準音声情報とを前記認識手段が順次比較することにより、類似する基準音声情報を認識し、類似する基準音声情報を記憶する類似ワード記憶手段とを備えた音声認識装置であって、前記数判断手段において、認識された音声に相当する基準音声情報の下位階層に属する音声認識辞書内の基準音声情報の語句の数が基準値を超えているか判断する判断手段を備え、前記所定数以上と判断されたとき、前記抽出手段が下位階層に属する前記音声認識辞書の一部の音声認識辞書を抽出して前記一覽記憶手段に記憶し、前記認識手段が前記一覽記憶手段に記憶されている基準音声情報との比較を終了した後、前記抽出手段が下位階層に属する前記音声認識辞書のうち未抽出の辞書を抽出し、前記一覽記憶手段が更新記憶するとともに、前記認識手段が前記一覽記憶手段内に更新記憶された辞書に属する基準音声情報と前記入力音声記憶手段に記憶されている音声とを順次比較し、類似する基準音声情報を認識し、類似ワード記憶手段が新たに認識された類似する前記基準音声情報を追加記憶することを特徴とする。

【 0 0 1 5 】

請求項 6 に記載の発明は、請求項 5 に記載の音声認識装置において、前記類似ワード記憶手段に記憶された全ての類似ワードにおける一つを前記認識手段が認識し、認識結果とすることを特徴とする。

【 0 0 1 6 】

請求項 7 に記載の発明は、請求項 5 に記載の音声認識装置において、前記一覽記憶手段に記憶されている基準音声情報の内、複数の類似する基準音声情報を前記類似ワード記憶手段に記憶し、前記類似ワード記憶手段に記憶された全ての類似する基準音声情報のなかから更に認識結果を選出する選出手段を備えることを特徴とする。

【 0 0 1 7 】

上述した課題を解決するために、請求項 8 に記載の発明は、複数の基準音声情報を有する複数の音声認識辞書を記憶した階層辞書部と、前記複数の音声認識辞

書のうち1の辞書を抽出する抽出手段と、抽出された前記辞書を記憶する一覧記憶手段と、音声を入力する音声入力手段と、入力された音声を記憶する入力音声記憶手段と、前記入力音声記憶手段に記憶されている音声と前記一覧記憶手段に記憶されている基準音声情報と順次比較し、類似する基準音声情報を認識する認識手段と、類似する前記基準音声情報を記憶する類似ワード記憶手段とを備えた音声認識装置であって、前記一覧記憶手段に記憶されている辞書に属する全ての基準音声情報と前記入力音声記憶手段に記憶されている音声との比較を前記認識手段が終了した後、前記抽出手段が音声認識辞書から未抽出の辞書を抽出し、前記一覧記憶手段が更新記憶するとともに、前記認識手段が前記一覧記憶手段に更新記憶された辞書に属する基準音声情報と前記入力音声記憶手段に記憶されている音声との比較をして類似する基準音声情報を認識し、前記類似ワード記憶手段が新たに認識した類似する基準音声情報を追加記憶することを特徴とする。

【0018】

請求項9に記載の発明は、請求項8に記載の音声認識装置において、前記類似ワード記憶手段に記憶されている複数の基準音声情報の中から、更に認識結果を選出する選出手段を備えたことを特徴とする。

【0019】

上記構成により、指定地点の絞込みをおこなう際に、ユーザが知らない条件の入力をシステムから求められた場合に、階層飛ばしを示す基準音声情報（「わからない」と発話）を入力ことにより、検索を継続することができ、操作性ならびに応答性の改善がはかれる。なお、この場合、絞込条件は予めシステムによって設定されていた数より少なくなってしまうため、最終的に名称を発話する際の認識対照語句の数が大きくなり、認識率が低下する可能性はあるが、検索を継続することが可能となり、操作性、応答性の面での効果が大きい。また、認識対象語句の増加により、メモリの容量が問題となるが、複数に分けて認識処理を行うことでこれを回避することができる。

【0020】

上述した課題を解決するために、請求項10に記載の発明は、階層構造をなす複数の音声認識辞書の中から基準音声情報を抽出し、抽出された基準音声情報と

入力された音声とを比較することにより音声を認識する音声認識方法であって、所定の音声認識辞書に階層飛ばしを示す基準音声情報が用意されており、前記階層飛ばしを示す基準音声情報に相当する音声の入力を認識した際に、比較対象となっていた基準音声情報の下位階層に属する音声認識辞書の一部を抽出して音声認識を行うことを特徴とする。

【 0 0 2 1 】

請求項 1 1 に記載の発明は、請求項 1 0 に記載の音声認識方法において、認識された基準音声情報の下位階層に属する音声認識辞書の中の基準音声情報の数を判断し、その数が基準値を超えていると判断されたとき、下位階層に属する前記音声認識辞書の一部を抽出して、比較することにより類似する基準音声情報を認識し、抽出された基準音声情報との比較が終了した後、下位階層に属する前記音声認識辞書の中から未抽出の音声認識辞書を抽出し、比較することにより類似する基準音声情報を認識し、複数の類似する前記基準音声情報の中から、入力された音声に相当する基準音声情報をさらに選出することを特徴とする。

【 0 0 2 2 】

上述した課題を解決するために、請求項 1 2 に記載の発明は、複数の基準音声情報を有する複数の音声認識辞書の中から 1 つの音声認識辞書を抽出し、抽出された音声認識辞書内の基準音声情報と入力された音声と比較し、抽出された前記音声認識辞書による基準音声情報との比較が終了した後、前記 1 つの音声認識辞書と異なる他の音声認識辞書を抽出し、抽出された音声認識辞書内の基準音声情報を、比較対象の基準音声情報として更新すると共に、更新された基準音声情報と前記入力された音声とを比較することにより、入力された前記音声を認識することを特徴とする。

【 0 0 2 3 】

【発明の実施の形態】

以下、本発明の実施例を示す図面を参照しつつ詳細に説明する。

【 0 0 2 4 】

カーナビゲーション装置に設けられた音声認識装置の一実施形態を示すブロック図を図 1 に示す。本発明の音声認識装置は、マイク 1 0 0 と、特徴量計算部 1

01と、認識部102と、RAM103と、認識辞書記憶部104と、認識辞書選択部105と、特徴量記憶部106と、認識結果記憶部107と、認識結果統合部108と、音声合成部109と、制御部110と、スピーカ111と、結果表示部112と、検索部113で構成される。

【0025】

マイク100はユーザが発話した音声を入力し、特徴量計算機101に出力する。

【0026】

特徴量計算部101は、マイク100によって取り込まれた音声信号をPCM (Pulse Code Modulation) データに変換し、当該PCMデータを音声認識に適した特徴量に変換し、認識部102と特徴利用記憶部106へ出力する。特徴量記憶部106は、計算された特徴量を記憶するとともに認識部102へ供給する。

【0027】

認識辞書記憶部104は、認識対象語句となる複数の基準音声情報を有する複数の音声認識辞書を互いに階層的に関連付けて記憶しているものである。辞書の種類として、複数の絞込条件毎に設けられた絞込条件辞書と、絞込条件の組み合わせによって分類される最終的な地点名称辞書とがある。地点名称辞書とは、例えば、”〇〇県××市に存在する教育施設に属する大学の全ての大学名称を示す基準音声情報を有する辞書”や”〇〇県××市に存在する病院に属する診療所の全ての診療所名称を示す基準音声情報を有する辞書”等、各地点に存在する具体的な各々の施設の名称を示す基準音声情報を記憶したものである。また、絞込条件辞書としては、地点検索を行うための都道府県名称等広い地域を示す地域名称を示す基準音声情報を記憶した都道府県名称辞書や、各都道府県毎に設けられ、各都道府県に属する市区町村名等狭い地域を示す地域名称を示す基準音声情報を記憶した市区町村名所辞書や、指定地点に存在する施設の種類等の大分類ジャンル名称を示す基準音声情報を記憶したジャンル名称辞書、及び、それぞれの大分類ジャンル毎に設けられ、各大分類ジャンルに属するサブジャンル名称を示す基準音声情報を記憶したサブジャンル名称辞書等がある。

【 0 0 2 8 】

なお、一般的に最下層の地点名称の総数は極めて多く、音声認識辞書用の R A M の容量及び認識率の点から、一度に認識する数としては実用的でないため、各地点名称辞書の地点名称の数（サイズ）がそれぞれ利用可能な R A M 1 0 3 の容量によって決定される基準数未満となるように、上述のように音声認識辞書が階層構造をなすとともに、複数の絞込条件の組合わせ毎に地点名称を分類し、分類毎に地点名称辞書が設けられている。

【 0 0 2 9 】

認識辞書選択部 1 0 5 は、制御部 1 1 0 からの認識対象とする音声認識辞書の抽出等の指令に応じて、認識辞書記憶部 1 0 4 の中から認識対象とする音声認識辞書を選択して抽出し、R A M 1 0 3 に供給する。R A M 1 0 3 は、音声認識辞書が供給される毎に、認識対象となる基準音声情報を供給された音声認識辞書のものに更新記憶する。

【 0 0 3 0 】

認識部 1 0 2 は、入力音声に変換された特徴量もしくは特徴量記憶部 1 0 6 に記憶された入力音声に変換された特徴量と、R A M 1 0 3 にロードされた音声認識辞書中の基準音声情報との類似度を計算し、類似度の高い基準音声情報及びその類似度（スコア）を認識結果として、認識結果記憶部 1 0 7 及び制御部 1 1 0 に出力する。

認識結果記憶部 1 0 7 には、認識部 1 0 2 により認識された認識結果（絞込条件または地点名称）もしくは制御部 1 1 0 から供給された認識結果を記憶するとともに、認識結果統合部 1 0 8 及び制御部 1 1 0 に出力する。認識結果統合部 1 0 8 は、認識結果記憶部 1 0 7 に複数の地点名称が認識結果として記憶されている場合に、類似度の高い方から K 個を決定し新たな認識結果として、制御部 1 1 0 へ供給する。そして、制御部 1 1 0 は、認識結果統合部 1 0 8 から供給された新たな認識結果を、第 2 認識結果として記憶更新すべく認識結果記憶部 1 0 7 に出力する。

【 0 0 3 1 】

音声合成部 1 0 9 は、制御部 1 1 0 からの指令によって、ガイダンスメッセー

ジやエコーバック用の合成音声を作成し、スピーカ111に供給する。スピーカ111は音声合成部109から供給された音声出力する。

【0032】

検索部111は、図示せぬ地図データ等のデータベースを備えており、制御部110からの指令によって、音声認識によって最終的に指定された地点の所在地図や住所や電話番号やサービス内容等の詳細な施設情報をデータベースから検索する。結果表示部112は、音声操作を行う際の認識結果、認識対象語句、ガイダンスメッセージ、エコーバック等とともに、検索部111によって検索された詳細な施設情報を表示するディスプレイである。

【0033】

制御部110は、上述した各構成から出力される出力結果に応じて各構成を制御するものである。即ち、制御部110は、音声操作による地点の指定がなされる際に、まず、ジャンル名称辞書を認識辞書選択部105が認識辞書記憶部104から取り出し、RAM103に認識対象となる基準音声情報として設定するように制御する。更に、認識部102から得られる認識結果や認識結果記憶部107に記憶された認識結果（絞込条件）を基に、適切な音声認識辞書を抽出すべく認識辞書選択部105への指示を行うとともに、音声合成部109へガイダンスメッセージの作成を指示する。

【0034】

また、認識結果統合部108から供給された新たな認識結果を、現在の認識結果として記憶更新すべく認識結果記憶部107に出力する。更に、最終的な認識結果（地点名称）を受けて、合成音声による認識結果のエコーバック、結果表示部112への結果表示、検索部113への検索指示等を行う。制御部110の動作の詳細についてはフローチャートを用いて後述する。

【0035】

ここで、認識辞書記憶部104に記憶されている複数の音声認識辞書が、互いに関連付けられて階層構造を形成している様を図2乃至図4を用いて説明する。

【0036】

なお、図2乃至図4は音声認識辞書の具体例の一部のみが示されている。まず

、最上位の第1階層の辞書として、階層飛ばし語句としての「わからない」及び「駅名」、「病院」、「宿泊施設」などのジャンル名称を示す基準音声情報を有するジャンル名称辞書（図2の200、図3の300、図4の400）が設けられている。

【0037】

第1階層に続く第2階層の辞書として、駅名、病院、宿泊施設などの各ジャンル名称に属するサブジャンル名称を示す基準音声情報を有する、サブジャンル名称辞書（図2の201、図3の302乃至305、図4の402乃至405）が、ジャンル名称毎に設けられている。また、サブジャンル名称を示す基準音声情報としては、例えば、病院のサブジャンル名称辞書では、診療所、内科、外科等に相当するサブジャンル名称を示す基準音声情報、及び階層飛ばし語句としての「わからない」を示す基準音声情報がある。

【0038】

更に、第2階層のサブジャンル名称辞書に続く第3階層の辞書として、日本全国の都道府県名称を示す基準音声情報及び階層飛ばし語句としての「わからない」を示す基準音声情報を有する都道府県名称辞書（図2の202、図3の306、図4の406）が設けられている。

【0039】

また、第3階層の都道府県名称辞書に続く第4階層の辞書として、各都道府県に存在する市区町村名称を示す基準音声情報及び階層飛ばし語句としての「わからない」を示す基準音声情報を有する市区町村名称辞書（図2の203、図3の308乃至311、図4の408乃至411）が、各都道府県名称毎に設けられている。

【0040】

そして、最下位の第5階層の辞書として、具体的な地点を示す該地点に存在する施設の固有名称などの地点名称（目的語）を示す基準音声情報を有する地点名称辞書（図2の204乃至210、図3の312乃至319、図413乃至420）が、第1階層から第4階層の絞込条件の組合わせ毎に設けられている。

【0041】

以下に順次条件の絞込みを行い、地点の指定を行う検索手順の例について説明する。

〔例 1〕

本例は、システムから催促される絞込条件をユーザが全て把握している場合の検索手順の例であり、ここでは埼玉県川越市の病院（診療所）の栗田医院を指定する例を、図 2 を用いて説明する。なお、図 2 における太線の矢印にて検索手順を示す。

【0042】

はじめに対象地点に存在する施設の種別（以下ジャンルと称する）での絞り込みを行う。”

ジャンル名称辞書” 200 を用意するとともに、(1)「ジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してジャンル名称の発話を促す。これに対してユーザがジャンル名称として例えば (2)「病院」と発話すると、音声認識システムが”ジャンル名称辞書” 200 を認識対象として発話を認識する。

【0043】

更なる絞込みを行うため、認識対象の”ジャンル名称辞書” 200 を”病院のサブジャンル名称辞書” 201 に入れ替えるとともに、(3)「次のジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してサブジャンル名称の発話を促す。これに対してユーザがサブジャンル名称として例えば (4)「診療所」と発話すると、音声認識システムが”病院のサブジャンル名称辞書” 201 を認識対象として発話を認識する。

【0044】

サブジャンルが確定すると、更に地域での絞込みを行う。認識対象の”病院のサブジャンル名称辞書” 201 を”都道府県名称辞書” 202 に入れ替えるとともに、(5)「都道府県名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して都道府県名称の発話を促す。これに対してユーザが都道府県名称として例えば (6)「埼玉県」と発話すると、音声認識システムが”都道府県名称辞書” 202 を認識対象として発話を認識する。

【0045】

都道府県が確定すると、更に詳細な地域での絞込みを行う。認識対象の”都道府県名称辞書”202を”埼玉県の市町村名称辞書”203に入れ替えるとともに、(7)「市区町村名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して市区町村名称の発話を促す。これに対してユーザが市区町村名称として例えば(8)「川越市」と発話すると、音声認識システムが”埼玉県の市町村名称辞書”203を認識対象として発話を認識する。

【0046】

ここで、システム側は認識対象の”埼玉県の市町村名称辞書”203を”埼玉県の川越市の診療所の地点名称辞書”204に入れ替えるとともに、(9)「名称をどうぞ」とガイダンスメッセージを出力し、指定地点の具体的名称の発話を促す。これに対してユーザが地点名称として(10)「栗田医院」と発話すると、音声認識システムが”埼玉県の川越市の病院の診療所の地点名称辞書”204を認識対象として発話を認識する。

【0047】

〔例2〕

本例は、システムから催促される絞込条件のうち対象地点が存在する市区町村名称をユーザが把握していない場合の検策手順の例であり、ここでは埼玉県の病院（診療所）の小林医院を指定する例を示す。なお、本例の検索手順を図3（a）における太線の矢印にて示す。

【0048】

はじめに、対象地点に存在する施設の種別（以下ジャンルと称する）での絞込みを行う。”ジャンル名称辞書”300を用意するとともに、(1)「ジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してジャンル名称の発話を促す。これに対してユーザがジャンル名称として例えば(2)「病院」と発話すると、音声認識が”ジャンル名称辞書”300を認識対象として発話を認識する。

【0049】

更なる絞込みを行うため、認識対象の”ジャンル名称辞書”300を”病院の

サブジャンル名称辞書” 303 に入れ替えるとともに、(3)「次のジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してサブジャンル名称の発話を促す。これに対してユーザがサブジャンル名称として例えば(4)「診療所」と発話すると、音声認識システムが” 病院のサブジャンル名称辞書” 303 を認識対象として発話を認識する。

【0050】

サブジャンルが確定すると、更に地域での絞込みを行う。認識対象の” 病院のサブジャンル名称辞書” 303 を” 都道府県名称辞書” 306 に入れ替えるとともに、(5)「都道府県名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して都道府県名称の発話を促す。これに対してユーザが都道府県名称として例えば(6)「埼玉県」と発話すると、音声認識システムが” 都道府県名称辞書” 306 を認識対象として発話を認識する。

【0051】

都道府県が確定すると、更に詳細な地域での絞込みを行う。認識対象の” 都道府県名称辞書” 306 を” 埼玉県の市町村名称辞書” 309 に入れ替えるとともに、(7)「市区町村名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して市区町村名称の発話を促す。これに対してユーザが市区町村名称を把握しておらず(8)「わからない」と階層飛ばし語句を発話すると、” 埼玉県の市町村名称辞書” 309 を認識対象として発話を認識する。

【0052】

第4階層にて階層飛ばし語句が発話された場合は、システム側は、第4階層の” 埼玉県の市町村名称辞書” 309 内の市町村名称の入力を催促せずに第4階層の辞書の絞込条件の入力を飛ばし、第4階層にて” 埼玉県の市町村名称辞書” 309 内の全ての市町村名称を絞込条件として入力されたものとし、第5階層の辞書として” 埼玉県の全ての市町村の病院の診療所の辞書” 313 乃至316 を抽出してまとめて” 埼玉県の病院の診療所の辞書” 312 を作成し、認識対象の” 埼玉県の市町村名称辞書” 309 を” 埼玉県の病院の診療所の辞書” 312 に入れ替えるとともに、(9)「名称をどうぞ」とガイダンスメッセージを出力し、指定地点の具体名称の発話を促す。これに対してユーザが地点名称として(10)「

小林医院」と発話すると、音声認識システムが” 埼玉県 of 病院 of 診療所 of 辞書” 312 を認識対象として発話を認識する。

【0053】

尚、上述の図3（a）は、ある階層での絞込条件の入力の際に飛ばし語句を発話すると、その階層での絞込条件の入力を飛ばして、すぐ下位の階層に進み絞込条件の入力を催促する例である。しかしながら、ジャンル名称が不明な時はサブジャンル名称も不明である可能性が高く、また、都道府県名称が不明な時は市区町村名称も不明な可能性が高いため、図3（b）に示すように、所定の階層にて飛ばし語句を発話した場合すぐ下位の階層へ進むのではなく、階層飛ばし語句の階層に応じて、2つ下の階層へ進むなどするように飛ばし先を設定することも考えられる。

【0054】

〔例3〕

本例は、システムから催促される絞込条件のうち対象地点に存在する施設のサブジャンルをユーザが把握していない場合の検索手順の例であり、ここでは埼玉県川越市の病院の斉藤医院を指定する例を示す。なお、本例の検索手順を図4における太線の矢印にて示す。

【0055】

はじめに対象地点に存在する施設の種別（以下ジャンルと称する）での絞込みを行う。”ジャンル名称辞書”400を用意するとともに、（1）「ジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してジャンル名称の発話を促す。これに対してユーザがジャンル名称として例えば（2）「病院」と発話すると、音声認識システムが”ジャンル名称辞書”400を認識対象として発話を認識する。

【0056】

更なる絞込みを行うため、認識対象の”ジャンル名称辞書”400を”病院のサブジャンル名称辞書”403に入れ替えるとともに、（3）「次のジャンルの名称をどうぞ」とガイダンスメッセージを出力し、ユーザに対してサブジャンル名称の発話を促す。これに対してユーザがサブジャンル名称を把握しておらず（4

)「わからない」と階層飛ばし語句を発話すると、音声認識システムが”病院のサブジャンル名称辞書”403を認識対象として発話を認識する。

【0057】

第2階層にて階層飛ばし語句が発話された場合は、システム側は、第2階層の”病院のサブジャンル名称辞書”403内のサブジャンル名称の入力を催促せずに第2階層の辞書の絞込条件の入力を飛ばし、第2階層にて”病院のサブジャンル名称辞書”403内の全てのサブジャンル名称を絞込条件として入力されたものとして、第3階層の認識対象の辞書として”病院のサブジャンル名称辞書”403を”都道府県名称辞書”406に入れ替えるとともに、(5)「都道府県名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して都道府県名称の発話を促す。これに対してユーザが都道府県名称として例えば(6)「埼玉県」と発話すると、音声認識システムが”都道府県名称辞書”406を認識対象として発話を認識する。

【0058】

都道府県が確定すると、更に詳細な地域での絞込みを行う。認識対象の”都道府県名称辞書”406を”埼玉県の市町村名称辞書”409に入れ替えるとともに、(7)「市区町村名をどうぞ」とガイダンスメッセージを出力し、ユーザに対して市区町村名称の発話を促す。これに対してユーザが市区町村名称として例えば(8)「川越市」と発話すると、音声認識システムが”埼玉県の市町村名称辞書”409を認識対象として発話を認識する。

【0059】

ここで、システム側は、”埼玉県の川越市の病院の辞書全て”417乃至420を抽出しまとめて”埼玉県川越市の病院の辞書”413を作成し、認識対象の”埼玉県の市町村名称辞書”309を”埼玉県川越市の病院の辞書”413に入れ替えるとともに、(9)「名称をどうぞ」とガイダンスメッセージを出力し、指定地点の具体的名称の発話を促す。これに対してユーザが地点名称として(10)「斉藤医院」と発話すると、音声認識システムが”埼玉県川越市の病院の辞書”413を認識対象として発話を認識する。

【0060】

図5乃至図7は、本発明実施形態の動作を説明するために引用したフローチャートである。

【0061】

以下、図5乃至図7に示すフローチャートを参照しながら、図1乃至図3に示す本発明実施形態の動作について詳細に説明する。

【0062】

図5において、まず制御部110は、図示せぬ発話ボタン等によりユーザからの音声入力による地点検索の検索開始要求がなされることを検出する。(ステップS500)。検出されなければ(ステップS500 NO)待機する。検出されたら(ステップS500 YES)認識結果記憶部107に記憶されている前回の絞込条件、すなわちジャンル名称、サブジャンル名称、都道府県名称、市区町村名称、及び指定地点の固有名称を全てクリアする(ステップS501)。認識辞書記憶部104に記憶された音声認識辞書の中からジャンル名称辞書を認識辞書選択部105に抽出させ、ジャンル名称辞書内の各ジャンル名称を認識対象語句とすべく各ジャンル名称を示す基準音声情報をRAM103にロードする(ステップS502)。

【0063】

制御部110は、ユーザにより発話された入力音声(ジャンル名または「わからない」)に対し、RAM103にロードされた音声認識辞書を対象として認識部102に認識処理を行わせ、認識結果を制御部110へ出力させる(ステップS503)。ステップS503で得られた認識結果が「わからない」等の階層飛ばし語句の場合、(ステップS504 YES)、ステップS505のジャンル名称による絞込条件の設定処理を飛ばして、ステップS506の処理へ進む。一方、ステップS503で得られた認識結果が、いずれのジャンル名称であった場合、(ステップS504 NO)、認識されたジャンル名称を絞込条件として認識結果記憶部107に記憶する(ステップS505)。

【0064】

続いて、制御部110は、認識辞書記憶部104に記憶されている音声認識辞書の中から、認識辞書選択部105に現在RAM103に認識対象語句として記

憶されているジャンル名称辞書に続く下位階層に属するサブジャンル名称辞書を抽出させ、抽出されたサブジャンル名称辞書内のサブジャンル名称を認識対象語句とすべく各サブジャンル名称を示す基準音声情報をRAM103にロードする（ステップS506）。なお、ここでロードされるサブジャンル名称は、ステップS503での認識結果が「わからない」等の階層飛ばし語句であった場合、ステップS502にてRAM103に認識対象語句としてセットされていたジャンル名辞書内の基準音声情報に属する下位階層に相当する全てのサブジャンル名辞書を選択するため、すべてのサブジャンル名称を認識対象語句としてRAM103にロードする。一方、ステップS503での認識結果が何らかのジャンル名称であった場合、認識されたジャンル名称に属するサブジャンル名称の音声認識辞書を選択して、選択されたサブジャンル名称辞書内のサブジャンル名称を対象語句としてRAM103にロードする。

【0065】

ユーザにより発話された入力音声（サブジャンル名または「わからない」）に対し、RAM103にロードされた音声認識辞書を対象として認識部102に認識処理を行わせ、認識結果を制御部110へ出力させる（ステップS507）。

【0066】

ステップS507にて得られた認識結果が「わからない」等の階層飛ばし語句の場合は、（ステップS508 YES）、ステップS509のサブジャンル名称による絞込条件の設定処理を飛ばして、ステップS510に進む。一方、ステップS507にて得られた認識結果が、いずれかのサブジャンル名称であった場合、（ステップS508 NO）、認識されたサブジャンル名称を絞込条件として認識結果記憶部107に設定する（ステップS509）。

【0067】

認識辞書記憶部104に記憶された音声認識辞書の中から、都道府県名称辞書を認識辞書選択部105に抽出させ、抽出させた都道府県名称辞書内の都道府県名称を認識対象語句とすべく各都道府県名称を示す基準音声情報をRAM103にロードする（ステップS510）。なお、ここでロードされる都道府県名称は、ステップS507での認識結果が「わからない」等の階層飛ばし語句あった場

合も、何らかのサブジャンル名であった場合も、都道府県名称辞書を選択して、選択された都道府県名称辞書内の都道府県名称を認識対象語句としてRAM103にロードする。

【0068】

ユーザにより発話された入力音声（都道府県名称または「わからない」）に対し、RAM103にロードされた音声認識辞書を対象として認識部102に認識処理を行わせ、認識結果を制御部110へ出力させる（ステップS511）。ステップS511で得られた認識結果が「わからない」等の階層飛ばし語句の場合は（ステップS512 YES）、ステップS513の都道府県名称による絞込条件の設定処理を飛ばして、ステップS514に進む。一方、ステップS511で得られた認識結果がいずれかの都道府県名称であった場合（ステップS512 NO）、認識された都道府県を絞込条件として認識結果記憶部107に設定する（ステップS513）。

【0069】

認識辞書記憶部104に記憶された音声認識辞書の中から、認識辞書選択部105に市区町村名称辞書を抽出させ、市区町村名称を認識対象語句とすべく各市区町村名称を示す基準音声情報をRAM103にロードする（ステップS514）。

【0070】

なお、ここでロードされる市区町村名称は、ステップS511での認識結果が「わからない」等の階層飛ばし語句あった場合、ステップS510にてセットされていた全国全ての都道府県名称辞書内の基準音声情報に属する下位階層に相当する全国全ての市区町村名称辞書を選択して、すべての市区町村名称を認識対象語句としてRAM103にロードする。一方、ステップS511での認識結果が何らかの都道府県名であった場合、認識された都道府県に存在する市区町村の音声認識辞書を抽出させ、抽出された市区町村名称辞書内の市区町村名称を認識対象語句としてRAM103にロードする。

【0071】

ユーザにより発話された入力音声（市区町村名称または「わからない」）に対

し、RAM103にロードされた音声認識辞書を対象として認識部102に認識処理を行わせ、認識結果を制御部110へ出力させる。(ステップS515)。

ステップS515で得られた認識結果が「わからない」等の階層飛ばし語句の場合は(ステップS516 YES)、ステップS517の市区町村名称による絞込条件の設定処理を飛ばしてステップS518へ進む。一方、ステップS515で得られた認識結果がいずれかの市区町村名であった場合(ステップS516 NO)、認識された市区町村を絞込条件として認識結果記憶部107に設定する(ステップS517)。

【0072】

認識辞書記憶部104に記憶された音声認識辞書でステップS505、S509、S513、S517の処理で認識結果記憶部107に記憶された絞込条件を満たす地点名称辞書内の基準音声情報の数(サイズ)の総和を計算する(ステップS518)。地点名称辞書のサイズの総和が、RAM103の容量に応じて設定された基準数を越える場合は(ステップS519 NO)、認識対象となる全ての地点名称辞書に対して複数回認識処理を行う(ステップS520)。地点名称辞書のサイズの総和がRAM103の容量以下の場合は(ステップS519 YES)、記憶された絞込条件を満たす全ての地点名称辞書内の地点名称を認識対象語句とすべく各地点名称を示す基準音声情報をRAM103にロードし(ステップS521)、通常の認識処理を行う(ステップS522)。そして、ステップS520またはステップS522で得られた認識結果である地点名称を出力させる(ステップS523)。

【0073】

尚、上述のフローチャートでは、絞込条件としてジャンル名称の入力を飛ばした場合、即ちステップS503で得られた認識結果が「わからない」等の階層飛ばし語句であった場合(ステップS504 YES)、ステップS505のジャンル名称による絞込条件の設定処理のみを飛ばして、ステップS506の処理へ進むようになっているが、上述の例に限らず、ジャンル名称がわからない場合サブジャンル名称もわからない可能性が高いため、サブジャンル名称の入力も飛ばして、ステップS510の処理へ進むようにしてもよい。

【 0 0 7 4 】

図 5 の示すステップ S 5 0 3、S 5 0 7、S 5 1 1、S 5 1 5、S 5 2 2 におけるユーザにより入力された音声に対する認識部 1 0 2 の各認識処理の詳細手順について、図 6 のフローチャートを用いて説明する。

【 0 0 7 5 】

図 6 において、マイク 1 0 0 からの音声の入力が開始されたか否かを判断する（ステップ S 6 0 0）。音声の入力の検出方法としては、特徴量計算部 1 0 1 にて、予め音圧レベルに関する閾値と基準時間とを記憶しておき、マイク 1 0 0 からの入力信号の音圧レベルと閾値とを比較し、入力信号が所定の閾値を越えた状態が、基準時間以上継続した場合に、音声の入力が開始されたとみなす等の方法が考えられる。

【 0 0 7 6 】

音声開始が検出されると、特徴量計算部 1 0 1 にて入力音声を音声認識に適した特徴量に変換させ（ステップ S 6 0 1）、特徴量記憶部 1 0 6 に記憶させるとともに、特徴量計算部 1 0 1 から認識部 1 0 2 に供給させる。認識部 1 0 2 にて、供給された特徴量と RAM 1 0 3 にロードされている各基準音声情報との類似度を計算する（ステップ S 6 0 2）。そして、音声の入力が終了されたか否かを判断する（ステップ S 6 0 3）。尚、音声の終了の検出方法としては、特徴量計算部 1 0 1 にて、予め音圧レベルに関する閾値と基準時間とを記憶しておき、マイク 1 0 0 からの入力信号の音圧レベルと閾値をそれぞれ比較し、入力信号が閾値以下の状態が所定時間以上続いた場合に、音声の入力が終了したとみなす等の方法が考えられる。

【 0 0 7 7 】

音声を終了していないと判断された場合（ステップ S 6 0 3 NO）、ステップ S 6 0 1 の処理へ戻る。一方、音声を終了したと判断された場合（ステップ S 6 0 3 YES）、ステップ S 6 0 2 にて求められた類似度の高い基準音声情報とその類似度とを対応づけて認識結果として、制御部 1 1 0 及び認識結果記憶部 1 0 7 に出力し、認識結果記憶部 1 0 7 に記憶させる（ステップ S 6 0 4）。

【 0 0 7 8 】

図5のフローチャートにて述べた、ステップS518にて求められた地点名称辞書のサイズの総和が、RAM103の容量に入りきらない場合（ステップS519 NO）に、ステップS520にて行われる複数回認識処理について、図7のフローチャートを用いて説明する。複数回認識処理とは、1回の入力音声に対して認識対象とする辞書（N個）を切り替えながら認識処理を繰り返し、それぞれの辞書に対する認識結果を統合し、最終的に全体としての認識結果を決定することである。

【0079】

図7において、認識辞書記憶部104に記憶された辞書でステップS505、S509、S513、S517の処理において認識結果記憶部107に記憶された絞込条件を満たす地点名称辞書の数（N）をカウントする（ステップS700）。続いて辞書番号 $n=1$ とする。（ステップS701）。ここで、絞込条件を満たす地点名称辞書のうち、管理番号が一番若い地点名称辞書を辞書番号1の地点名称辞書とし、認識辞書選択部105に辞書番号 $n(=1)$ 番目の地点名称辞書を認識辞書記憶部104から抽出させ、抽出した地点名称辞書の地点名称を認識対象語句とすべく各地点名称を示す基準音声情報をRAM103にロードする（ステップS702）。ここで、管理番号とは、認識辞書記憶部104に記憶された音声認識辞書各々に順番につけられた番号のことである。

【0080】

次に、マイク100から音声の入力が開始されたか否かを判断する（ステップS703）。音声の入力の検出方法としては、特徴量計算部101にて、予め音圧レベルに関する閾値と基準時間とを記憶しておき、マイク100からの入力信号の音圧レベルと閾値をそれぞれ比較し、入力信号が所定の閾値を越えた状態が、基準時間以上継続した場合に、音声の入力が開始されたとみなす等の方法が考えられる。

【0081】

音声開始が検出されると、特徴量計算部101にて入力音声を音声認識に適した特徴量に変換させ（ステップS704）、特徴量記憶部106に記憶させ（ステップS705）、特徴量記憶部106から認識部102に供給させる。認識

部102にて、供給された特徴量とRAM103にロードされている全ての基準音声情報との類似度を計算させる（ステップS706）。そして、音声の入力が終了したか否かを判断する（ステップS707）。

【0082】

尚、音声の終了の検出方法としては、特徴量抽出部101にて、予め音圧レベルに関する閾値と基準時間とを記憶しておき、マイク100からの入力信号の音圧レベルと閾値をそれぞれ比較し、入力信号が所定の閾値以下の状態が所定時間続いた場合に、音声の入力が終了したとみなす等の方法が考えられる。

【0083】

音声が終わっていないと判断された場合（ステップS707 NO）、ステップS704の処理へ戻る。一方、音声の終了したと判断された場合（ステップS707 YES）、ステップS706にて求められた類似度の高い順にK個の基準音声情報を、その類似度とを対応づけて辞書番号 $n=1$ の地点名称辞書の認識結果として、認識結果記憶部107に出力し、認識結果記憶部107に記憶させる（ステップS708）。なお、Kは1以上の整数であり、適宜システム設計者によって設定される値である。

【0084】

続いて辞書番号 $n=2$ とする（ステップS709）。辞書番号 n が、ステップS700にてカウントした認識対象辞書数（N）より大きいか否かが判断される（ステップS710）。辞書番号 n が認識対象辞書数（N）以下であれば（ステップS710 NO）、ステップS711の処理へ進む。絞込条件を満たす地点名称辞書のうち、管理番号が n 番目に若い地点名称辞書を辞書番号 $=n$ の地点名称辞書とし、認識辞書選択部105に辞書番号（ n ）の地点名称辞書を認識辞書記憶部104から抽出させ、抽出した地点名称辞書の地点名称を認識対照語句とすべく各地点名称を示す基準音声情報をRAM103にロードする（ステップS711）。

【0085】

入力音声の特徴量は既に特徴量記憶部106に記憶されているので、そこから認識部102へ供給し認識部102にて、供給された特徴量とRAM103にロ

ードされている全ての基準音声情報との類似度を計算させる（ステップ S 7 1 2）。ステップ S 7 1 2 にて求められた類似度の高い順に K 個の基準音声情報を、その類似度とを対応づけて辞書番号 n の地点名称辞書の認識結果として、認識結果記憶部 1 0 7 に出力し、認識結果記憶部 1 0 7 に記憶させる（ステップ S 7 1 3）。そして、辞書番号 n を $n + 1$ にインクリメントとする（ステップ S 7 1 4）。以下、ステップ S 7 1 0 にて辞書番号 n が認識対象辞書数 (N) を越えると判断されるまで、ステップ S 7 1 1 ~ ステップ S 7 1 4 の処理を繰り返す。

【 0 0 8 6 】

一方、辞書番号 n が認識対象辞書数 (N) より大きければ（ステップ S 7 1 0 Y E S）、ステップ S 7 1 5 の処理へ進む。ステップ S 7 1 5 では認識結果統合部 1 0 8 によって認識結果記憶部 1 0 7 に記憶された $K \times N$ 個の認識結果の中から類似度の高い順に第 K 個を第 2 認識結果として選択して、制御部 1 1 0 に出力させ、認識結果記憶手段 1 0 7 に更新記憶させる。尚、K が 1 の場合は、ステップ S 7 1 5 にて認識結果が 1 つ特定されるが、K が 2 以上の場合は、K 個の第 2 認識結果の中から、さらに 1 つを選択させるため、K 個の第 2 認識結果を制御部 1 1 0 に出力し、結果表示部 1 1 2 に K 個の地点名称を表示させ、図示せぬ操作釦にて選択させるようにする。もしくは、類似度が一番高いものを認識結果としてスピーカ 1 1 1 及び結果表示部 1 1 2 を用いてユーザに提示し、ユーザの N O 等の発話に応じて、続いて高い類似度のものを同様に提示し、ユーザが操作もしくは Y E S 等と発話されるまで順次提示して、認識結果から一つを決定するようにしてもよい。

【 0 0 8 7 】

なお、階層飛ばし語句としては、「わからない」という言葉は一例であり、システムが要求している情報をユーザが有していないことを表現する言葉であれば良く、例えば「パス」「次」等複数であっても良い。また、絞込の条件も「ジャンル名称」「サブジャンル名称」「都道府県名称」「市区町村名称」に限定されるものではなく、「地名名称」や「郵便番号」などでもよい。

【 0 0 8 8 】

【発明の効果】

以上説明のように本発明によれば、指定地点の絞込みをおこなう際に、ユーザが知らない条件の入力をシステムから求められた場合に、階層飛ばしを示す基準音声情報（「わからない」と発話）を入力ことにより、検索を継続することができ、操作性ならびに応答性の改善がはかれる。

【 0 0 8 9 】

なお、この場合、絞込条件は予めシステムによって設定されていた数より少なくなってしまうため、最終的に名称を発話する際の認識対照語句の数が大きくなり、認識率が低下する可能性はあるが、検索を継続することが可能となり、操作性、応答性の面での効果が大きい。また、認識対象語句の増加により、メモリの容量が問題となるが、複数に分けて認識処理を行うことでこれを回避することができる。

【図面の簡単な説明】

【図 1】 本発明における音声認識装置の実施形態を示すブロック図である。

【図 2】 本発明において使用される階層構造を持つ音声認識辞書の階層辞書ツリーの一例を示す図である。

【図 3】 本発明において使用される階層構造を持つ音声認識辞書の階層辞書ツリーの一例を示す図である。

【図 4】 本発明において使用される階層構造を持つ音声認識辞書の階層辞書ツリーの一例を示す図である。

【図 5】 本発明実施形態の音声認識処理による地点検索の動作を説明するために引用したフローチャートである。

【図 6】 本発明実施形態の音声認識処理動作を説明するために引用したフローチャートである。

【図 7】 本発明実施形態の複数回認識処理動作を説明するために引用したフローチャートである。

【符号の説明】

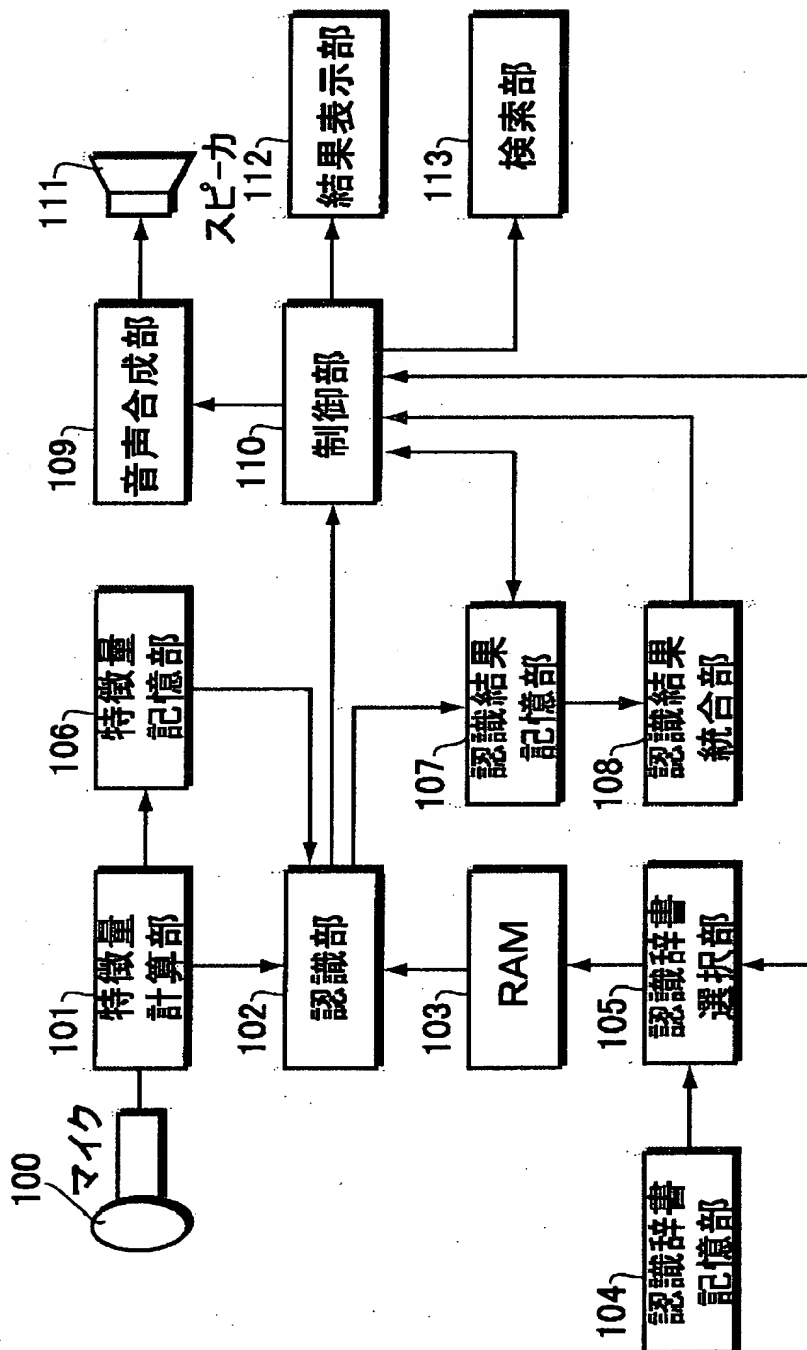
1 0 0 …マイク、1 0 1 …特徴量計算部、1 0 2 …認識部、1 0 3 …RAM、1 0 4 …認識辞書記憶部、1 0 5 …認識辞書選択部、1 0 6 …特徴量記憶部、1 0 7 …認識結果記憶部、1 0 8 …認識結果統合部、1 0 9 …音声合成部、1 1 0 …

制御部、 1 1 1 …スピーカ、 1 1 2 …結果表示部、 1 1 3 …検索部

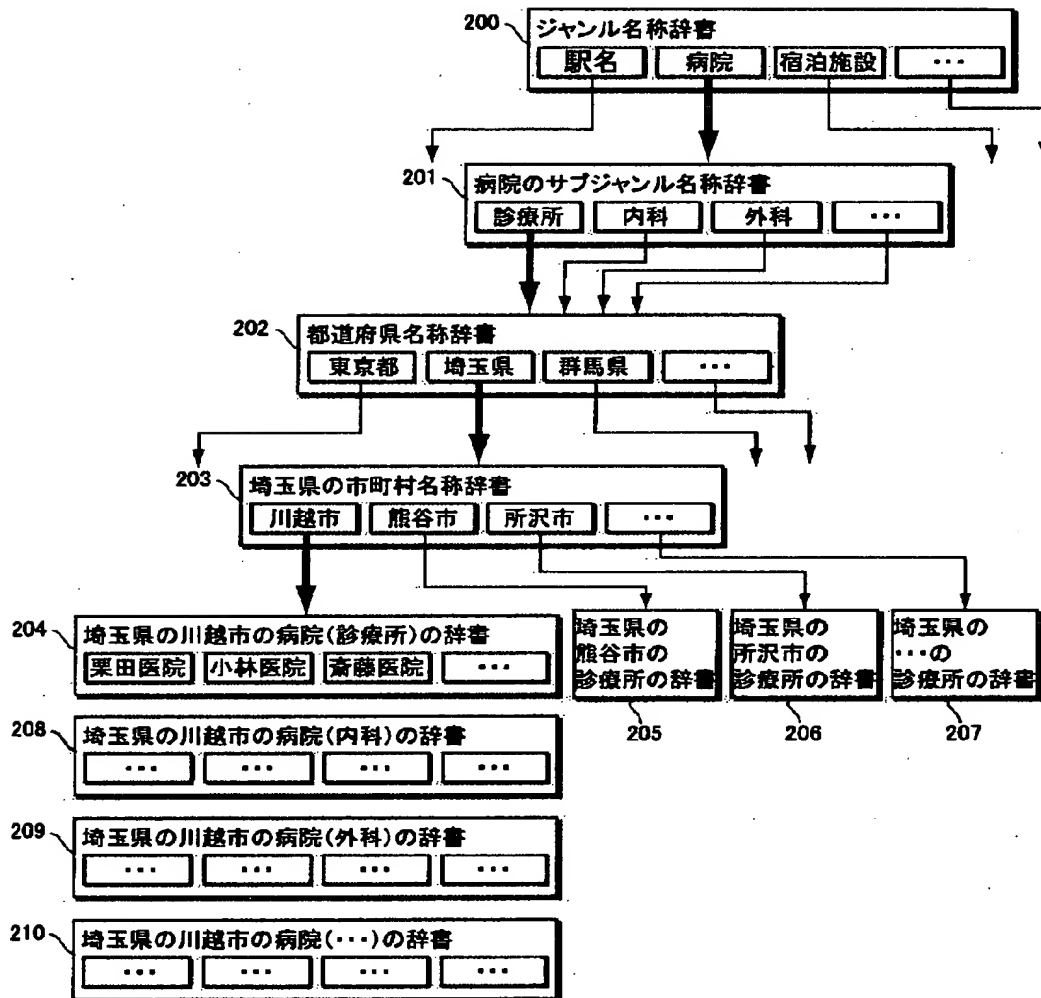
【書類名】

図面

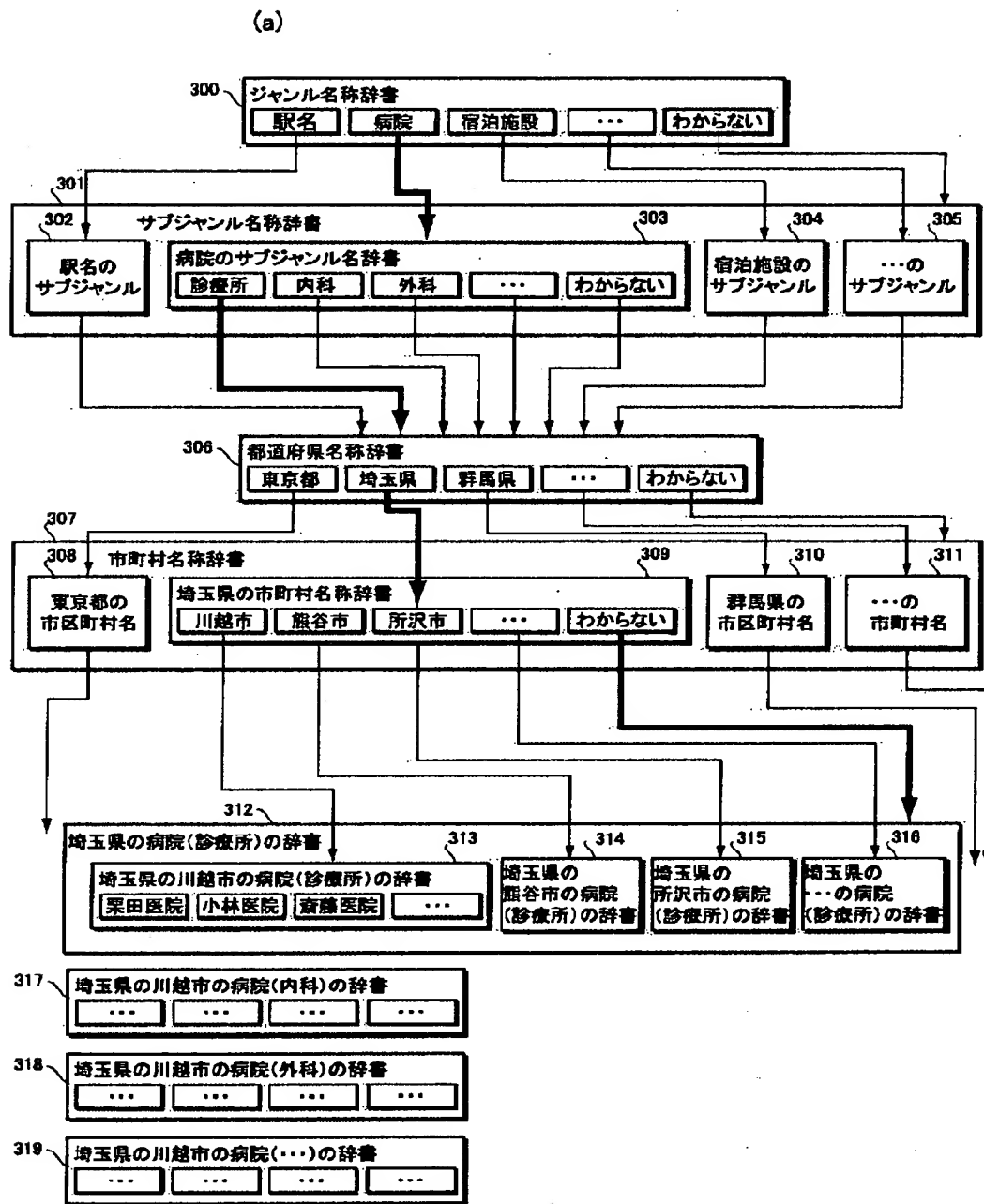
【図 1】



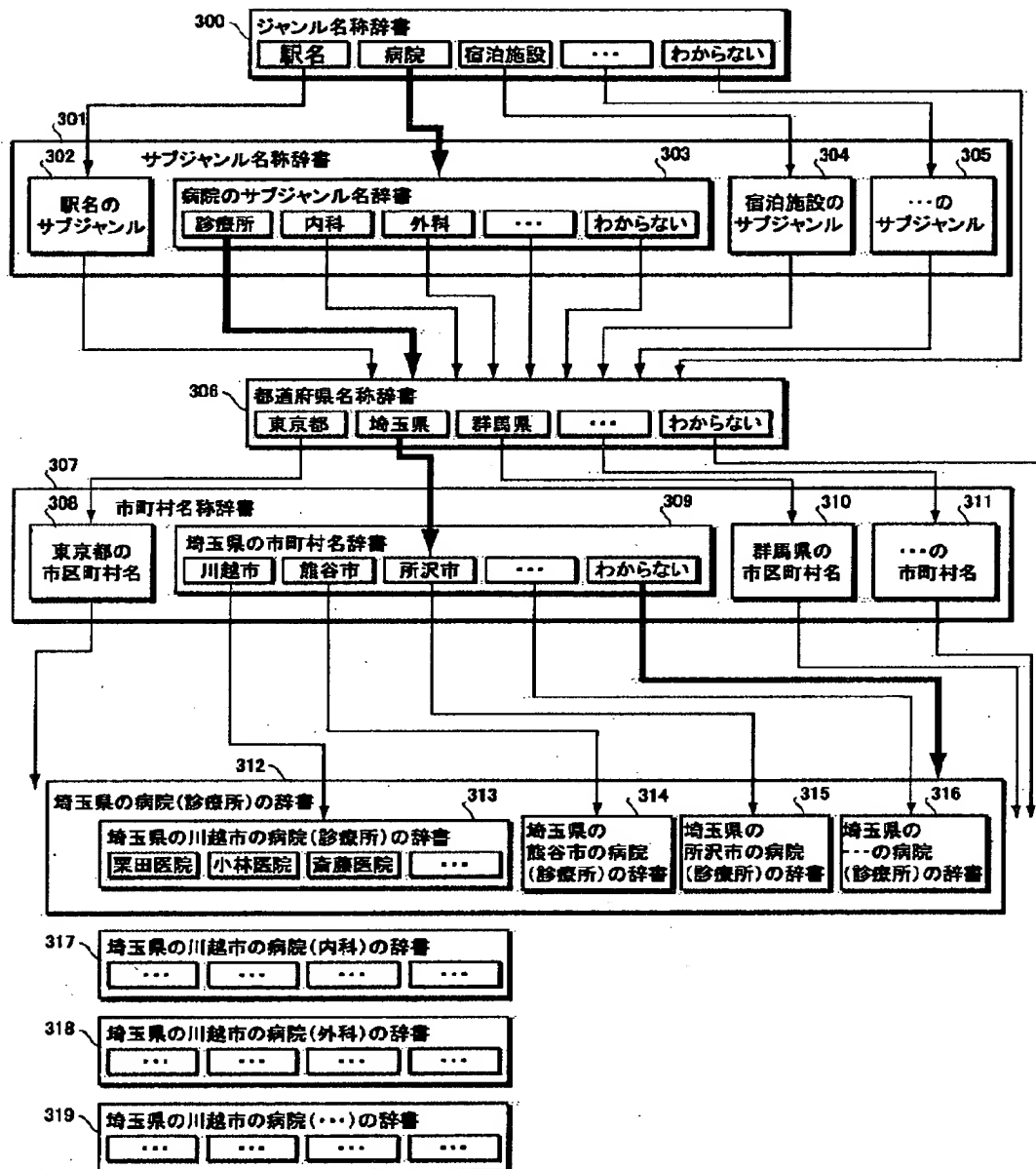
【図 2】



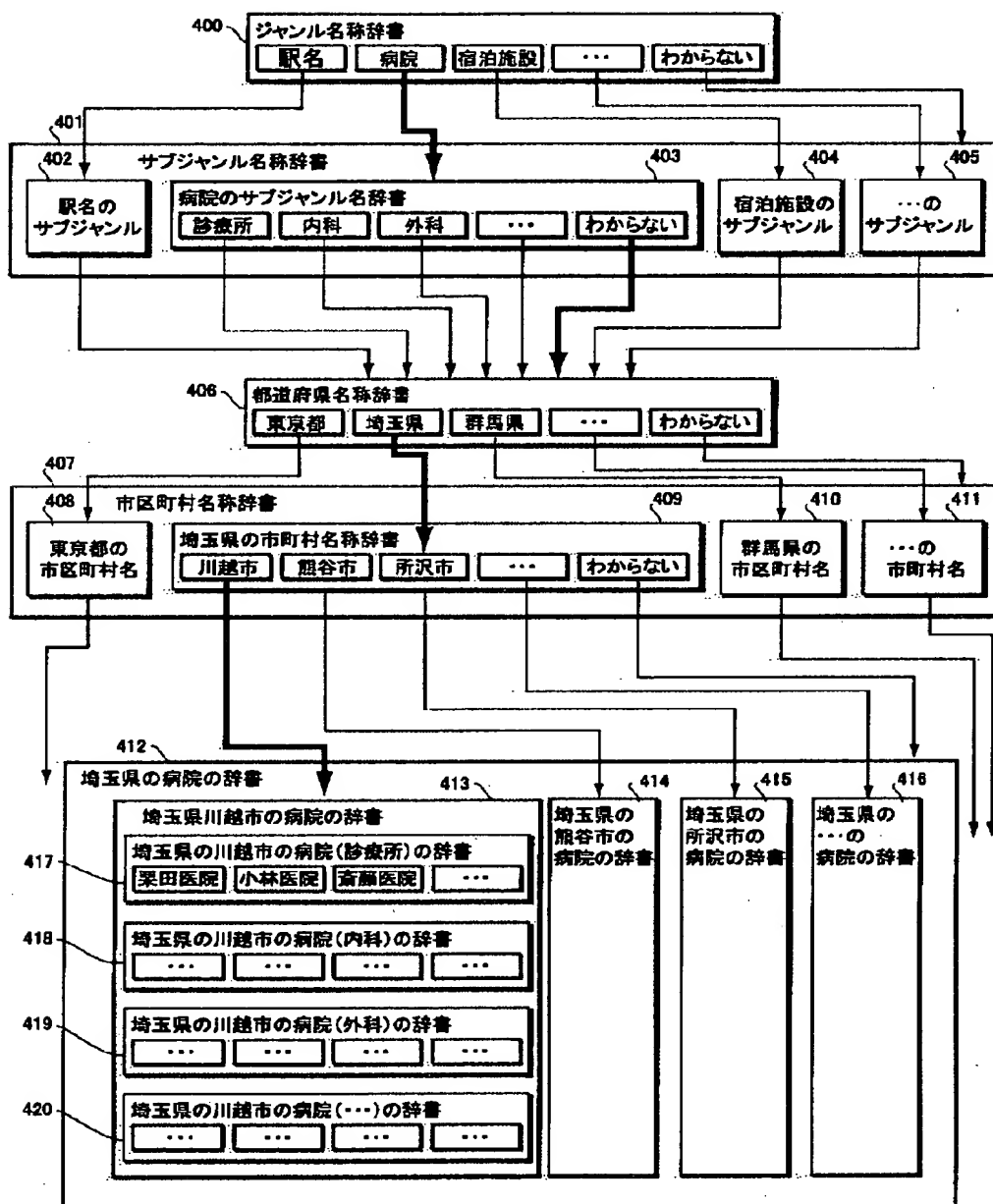
【図 3】



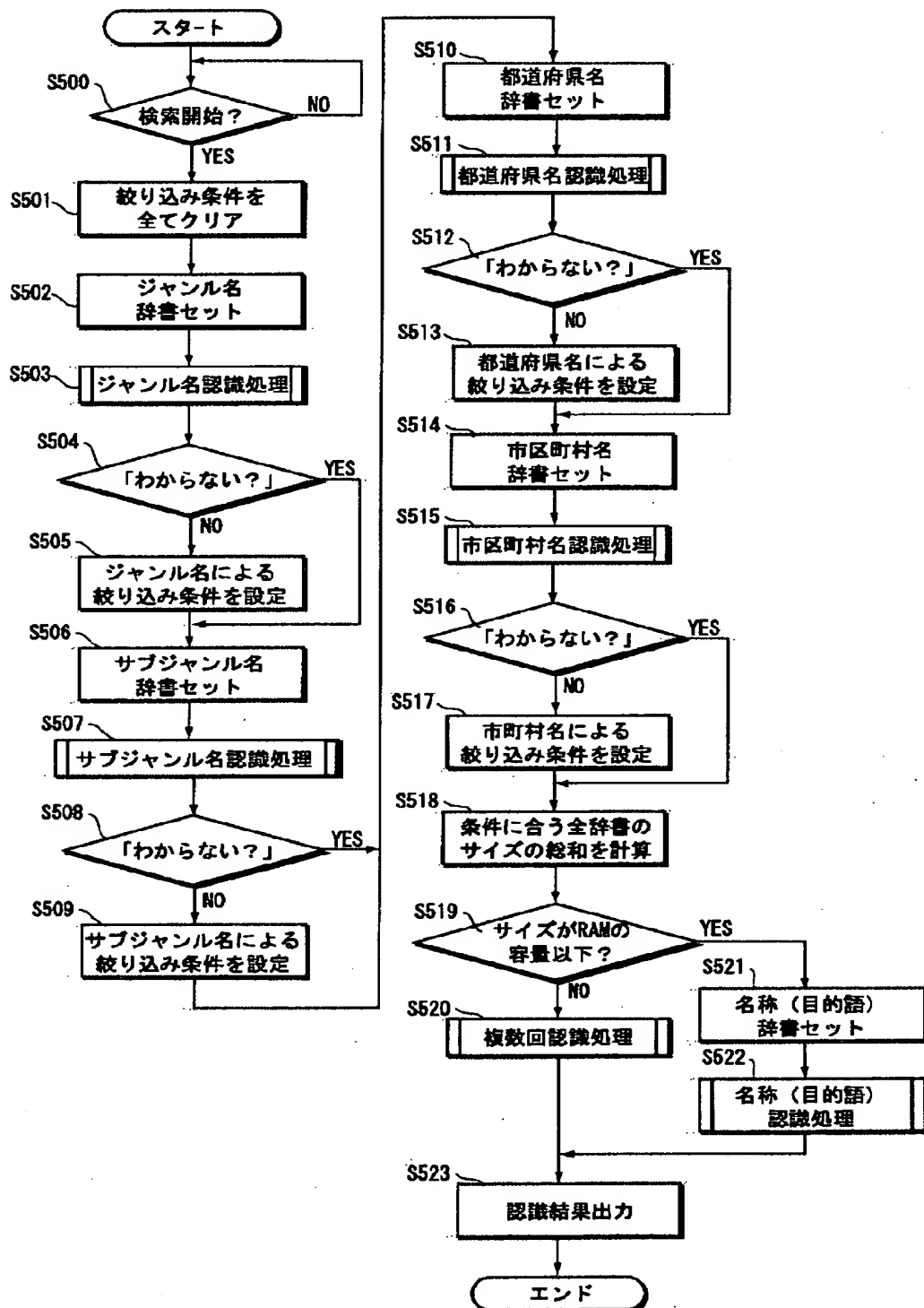
(b)



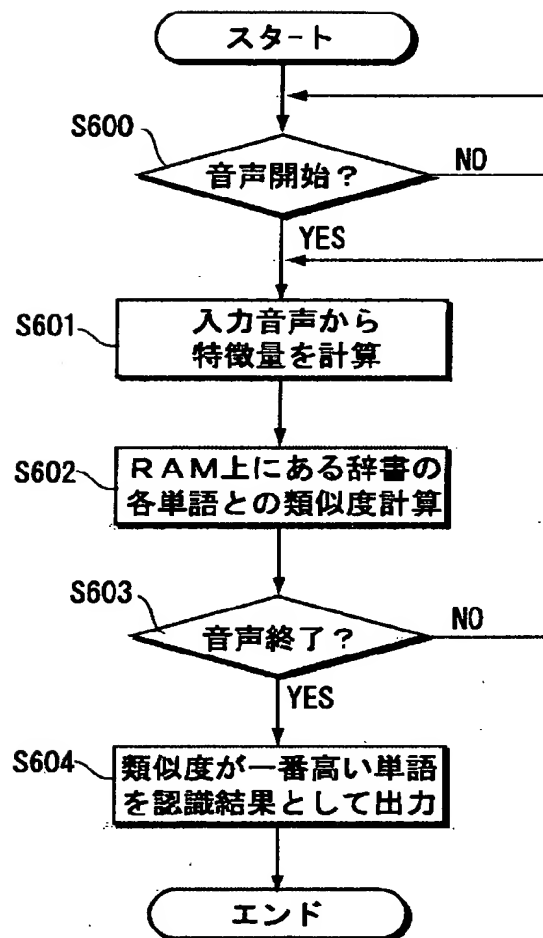
【図 4】



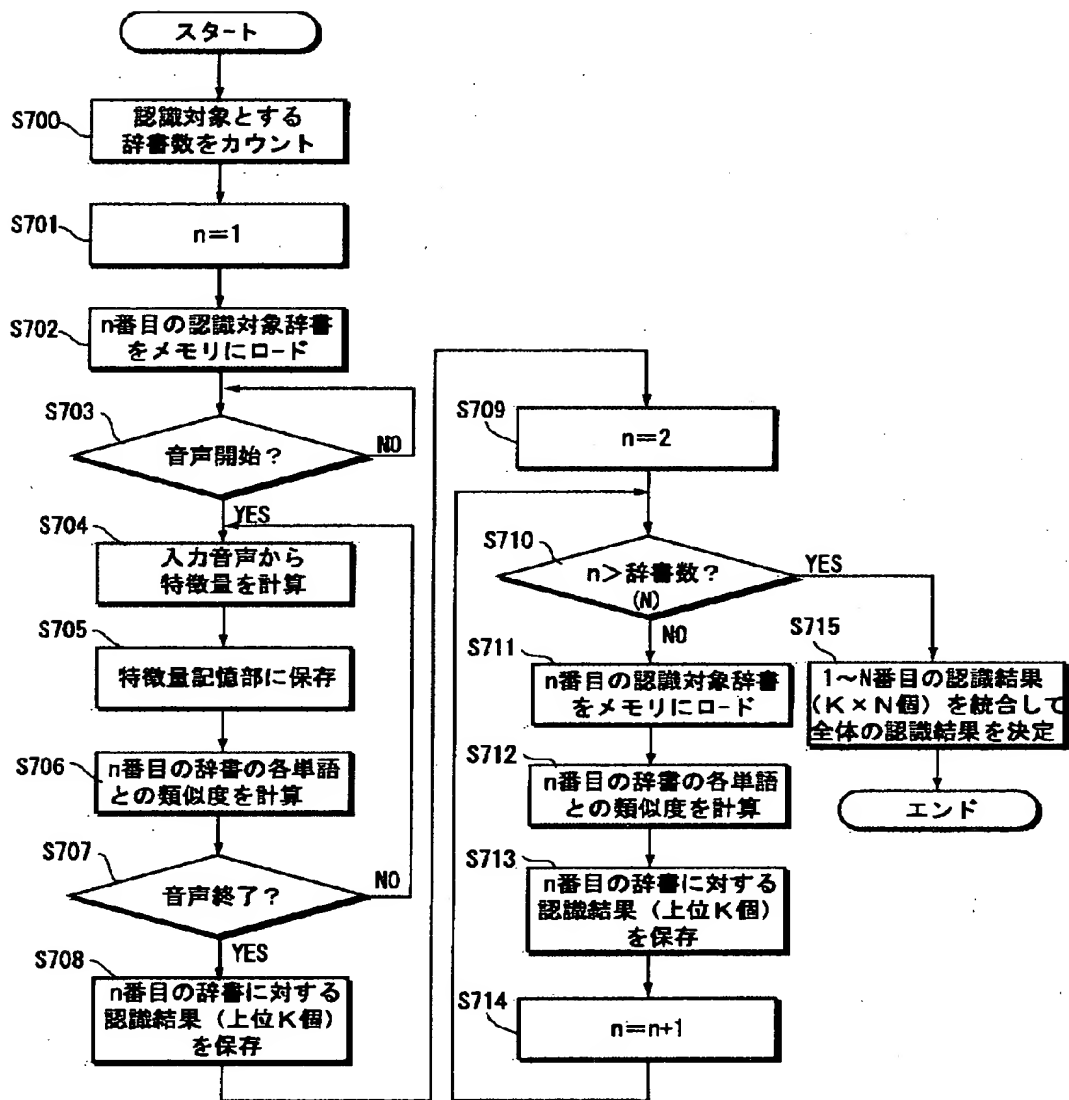
【図 5】



【図 6】



【図 7】



【書類名】 要約書

【要約】

【課題】 階層的に絞込条件を入力して地点指定を行う際に、システムにユーザが知らない絞込条件の入力が要求された場合でも、地点指定を行うための以降の継続するステップに進むことのできる仕組みを構築し、操作性、ならびに応答性の改善をはかる。

【解決手段】 階層構造をなす複数の音声認識辞書の中から基準音声情報を抽出し、抽出された基準音声情報と入力された音声とを比較することにより音声を認識する音声認識に関し、所定の音声認識辞書に階層飛ばしを示す基準音声情報を用意し、階層飛ばしを示す基準音声情報に相当する入力を認識した際に、比較対象となっていた基準音声情報の下位階層に属する音声認識辞書の一部を抽出して音声認識を行う。

【選択図】 図 5

出 願 人 履 歴 情 報

識別番号 [000005016]

| | |
|----------|-----------------|
| 1. 変更年月日 | 1990年 8月31日 |
| [変更理由] | 新規登録 |
| 住 所 | 東京都目黒区目黒1丁目4番1号 |
| 氏 名 | パイオニア株式会社 |